Data, Data, Everywhere—
How Shall We Live With It?

Cybersecurity Lessons
From Studying Big Data and Privacy

Marjory S. Blumenthal
RAND *(but not a RAND talk)*

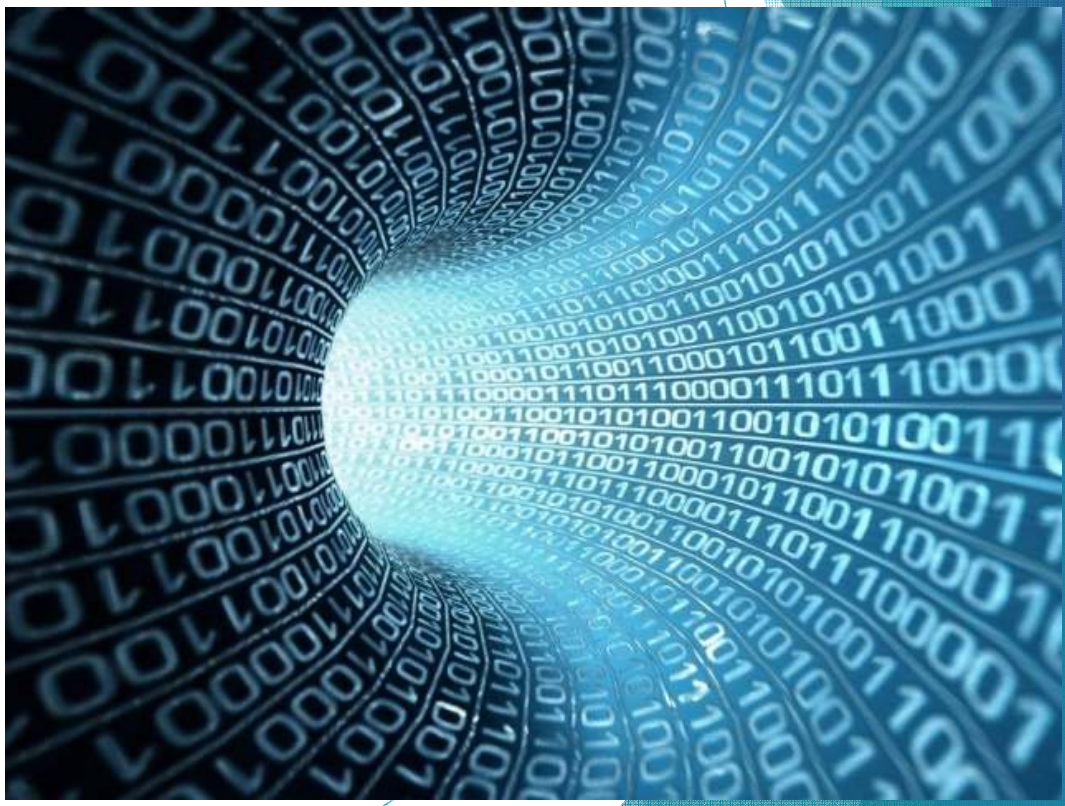
http://hdl.handle.net/2022/21733

Outline

- Why Big Data matters—a privacy-and-provenance perspective
- Through the lens of key domains
- Cybersecurity in/as context
- Moving forward

What Makes Big Data Special? Why Does It Challenge Privacy?

- Dimensions of big-ness
  - Quantity and variety of data (3 Vs)
  - Scale of analysis (“analytics”)
- Big data is dual use



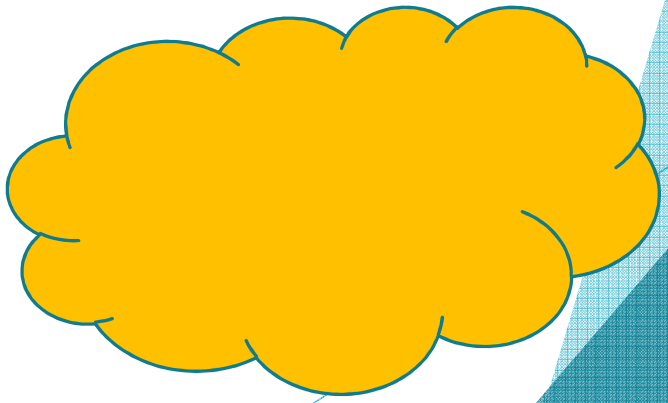
Feeding Those Vs and Privacy Concerns: People Emit Data Continuously . . .

Born digital

Born analog

Over-collection?

- Collection increasingly easy
- Cloud as dominant platform



Key Domains

Online Education Platforms

Source: College Informations, “Online Education Master Degree” http://collegeinformations.com/tag/online-education/

Health Care, Including mHealth

Source: UT School of Biomedical Informatics, “Emerging Trends in mHealth”
https://sbmi.uth.edu/blog/jan-15/01142015.htm

Technology in the Home

Source: Geekologie roomba-floor-plans-for-sale.jpg

What Is To Be Done?

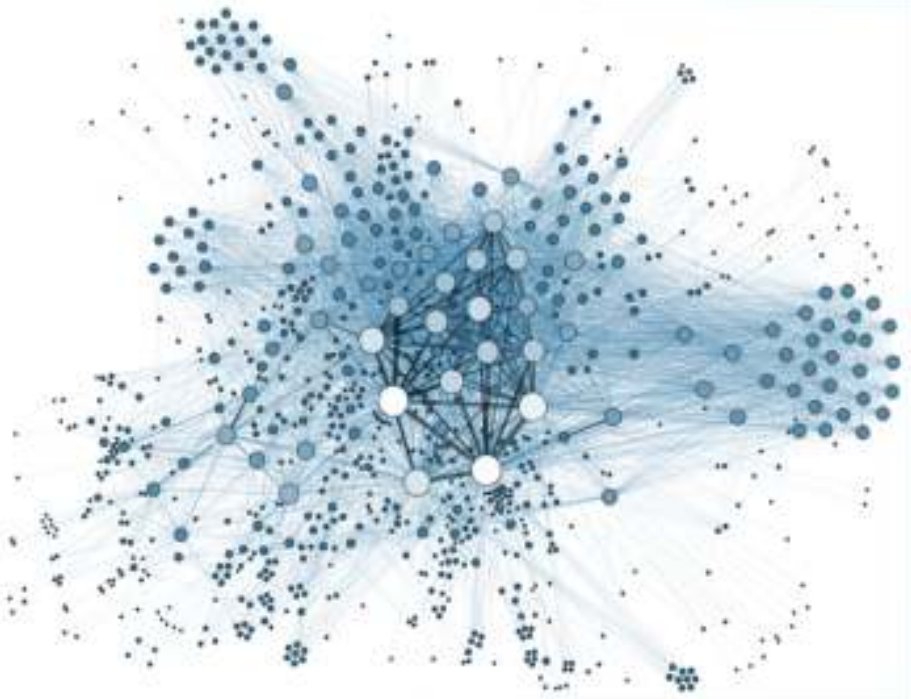
Cybersecurity: Necessary But Not Sufficient *to Protect Privacy*

- Cybersecurity technologies enforce policies…

- Cloud pros and cons

Analytics Can Violate Privacy W/O Security Violations

- Data fusion and integration
- Data mining and machine learning (“AI”)
- Techniques for classifying/finding relationships
- Image/speech recognition
- Social-network analysis



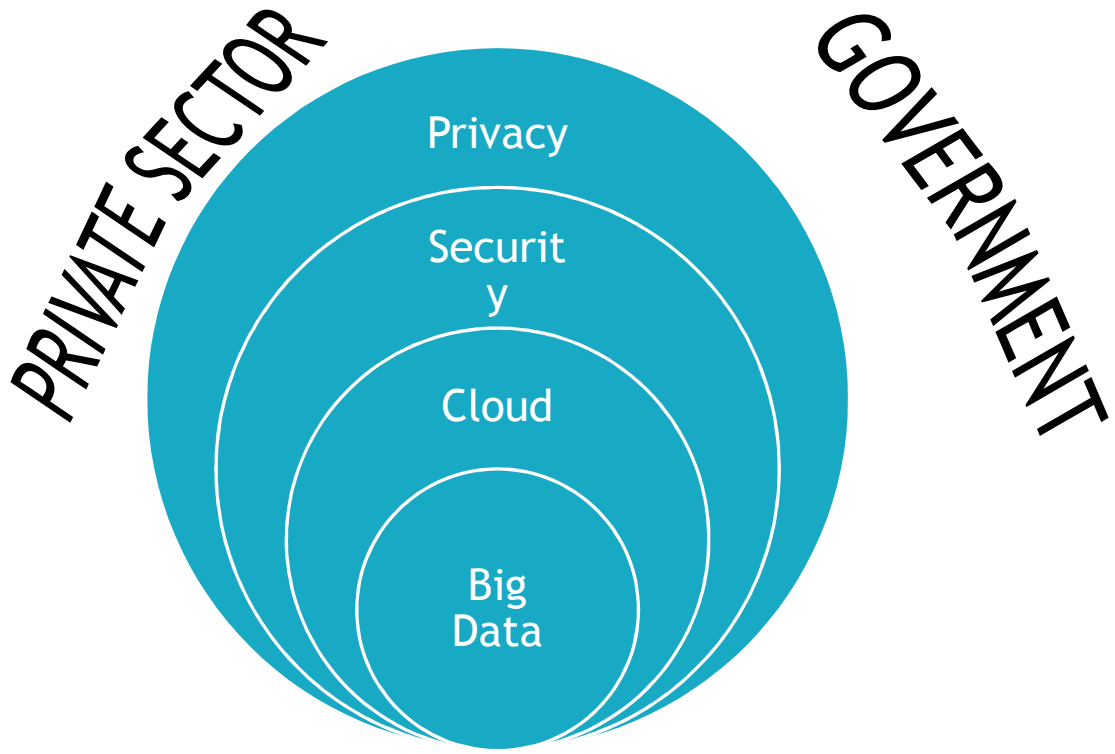
Cybersecurity Components

- Confidentiality
- Availability
- Integrity
  - *Fourth V: Veracity?*

Tools for Privacy Protection

- Cryptography and encryption
- Anonymization and de-identification
- Data deletion and ephemerality

A Systems Problem . . .



One View on Moving Forward: PCAST's Recommendations

1. Focus more on actual uses, less on collection
2. Design policy around outcomes, not specific technologies
3. Strengthen U.S. research in privacy-related technologies and relevant social science
4. Promote increased education and training for privacy protection
5. The U.S. Government should lead by example, sustain a leading position

Questions?

Thank You

marjory@rand.org

Podesta Recommendations, Aftermath

1. Advance the Consumer Privacy Bill of Rights
2. Pass National Data Breach Legislation
3. Extend Privacy Protections to Non-U.S. Persons
4. Ensure Data Collected on Students in School is used for Educational Purposes
5. Expand Technical Expertise to Stop Discrimination
6. Amend the Electronic Communications Privacy Act